



US005552281A

United States Patent [19]
Stashenko et al.

[11] Patent Number: **5,552,281**
[45] Date of Patent: **Sep. 3, 1996**

[54] HUMAN OSTEOCLAST-SPECIFIC AND -RELATED GENES

[75] Inventors: Philip Stashenko, Norfolk; Yi-Ping Li, Boston; Anne L. Wucherpfennig, Brookline, all of Mass.

[73] Assignee: Forsyth Dental Infirmary for Children, Boston, Mass.

[21] Appl. No.: 392,678

[22] Filed: Feb. 23, 1995

Related U.S. Application Data

[63] Continuation of Ser. No. 45,270, Apr. 6, 1993, abandoned.

[51] Int. Cl.⁶ C07H 21/04; C12N 5/10; C12N 15/70; C12Q 1/68

[52] U.S. Cl. 435/6; 435/69.1; 435/172.3; 435/252.3; 435/320.1; 536/23.1

[58] Field of Search 435/6, 320.1, 252.3, 435/69.1, 172.3; 536/23.1

[56] **References Cited**

PUBLICATIONS

Blair, Harry C., et al., "Extracellular-matrix degradation at acid pH. Avian osteoclast acid collagenase isolation and characterization", *Biochemical Journal* 290(3):873-884 (15 Mar. 1993).

Tezuka, Ken-Ichi, et al., "Identification of osteopontin in isolated rabbit osteoclasts", *Biochemical and Biophysical Research Communications* 186(2):914-916 (31 Jul. 1992). Tezuka, Ken-Ichi, et al., "Molecular cloning of a possible cysteine proteinase predominantly expressed in osteoclasts", *Journal of Biological Chemistry* 269(2):1106-1108, (14 Jan. 1994).

Horton, Michael A. et al., "Monoclonal Antibodies to Osteoclastomas (Giant Cell Bone Tumors): Definition of Osteoclast-specific Cellular Antigens," *Cancer Research* 45, 5663-5669 (Nov. 1985).

Davies, John et al., "The Osteoclast Functional Antigen, Implicated in the Regulation of Bone Resorption, Is Biochemically Related to the Vitronectin Receptor," *The Journal of Cell Biology* 109, 1817-1826 (Oct. 1989).

Hayman, Alison, R. et al., "Purification and characterization of a tartrate-resistant acid phosphatase from human osteoclastomas," *Biochem. J.* 261, 601-609 (1989).

Sandberg, M. et al., "Localization of the Expression of Types I, III, and IV Collagen, TGF- β 1 and c-fos Genes in Developing Human Calvarial Bones," *Developmental Biology* 130, 324-334 (1988).

Sandberg, M. et al., "Enhanced expression of TGF- β and c-fos mRNAs in the growth plates of developing human long bones," *Development* 102, 461-470 (1988).

Ek-Rylander, Barbro et al., "Cloning, Sequence, and Developmental Expression of a Type 5, Tartrate-resistant, Acid Phosphatase of Rat Bone," *The Journal of Biological Chemistry* 266(36), 24684-24689 (Dec. 25, 1991).

GenBank/EMBL Sequence Search Printout, pp. 1-19 (Jun. 24, 1993).

Primary Examiner—W. Gary Jones

Assistant Examiner—Paul B. Tran

Attorney, Agent, or Firm—Hamilton, Brook, Smith & Reynolds, P.C.

[57] **ABSTRACT**

The present invention relates to purified DNA sequences encoding all or a portion of an osteoclast-specific or -related gene products and a method for identifying such sequences. The invention also relates to antibodies directed against an osteoclast-specific or -related gene product. Also claimed are DNA constructs capable of replicating DNA encoding all or a portion of an osteoclast-specific or -related gene product, and DNA constructs capable of directing expression in a host cell of an osteoclast-specific or -related gene product.

5 Claims, 1 Drawing Sheet

```

1 AGACACCTCT GGCCTTCAACCA TCGAGCTTCG CGCCCGCTG GTCCTGGCTG
61 GCGCTTGCTG TTTCCTGCC CGACAGACG CGAAGTTCTG CTTCGTTGCTG TCCCTGAGA
121 CCTTCAGCC AAATCTCCG CGAGGAGCT CGAGAGAA TACTCTGGT GCTCTGTTA
181 CCTGGGGGTG CGACAGACG GTGGAGCTG GAAATCTTG CGGGCTGG CGCTGCTG
241 CGAGAGAA CGTCGCTGC CGAGAGAA CGGGCTGG CGCTGCTG AGGGCGAT
301 CGGGAGCGA CGTGGGGCTG TCTCAGCTG GTCGGGATG CAAAGCTTG AGGGCGAT
361 CAAGAGCGA CGCACAGACA TCGCTTATG GATCCAGAAC TACTGGCGG AGTTGGCGG
421 GGCGGGTGGT GAGGAGCTG TTTCCTGCC CGTCGCGCTG TCGGGCTGG TGAACGGCT
481 GAGCTCTGC CGCTGCTGCTG CGCTGCTG CGAGAGCTG ATGGGGTTG GTGGGGGGG
541 CGGGAGCGA CGTGGGGCTG CGGGCTGG CGCTGCTG CGGGCTGG CGCTGCTG
601 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
661 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
721 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
781 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
841 GAGCTCTGC CGCTGCTGCTG CGCTGCTG CGAGAGCTG ATGGGGTTG GTGGGGGGG
901 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
961 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
1021 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
1081 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
1141 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
1201 TTTCCTGGGGGGG ATGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG ATGGGGGGG
1261 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
1321 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
1381 AGCTCTGC CGAGAGACG CGAGAGAA CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
1441 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
1501 AGCTCTGC CGAGAGACG CGAGAGAA CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
1561 TTTCCTGGGGGG ATGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
1621 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
1681 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
1741 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
1801 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
1861 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
1921 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
1981 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
2041 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
2101 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
2161 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
2221 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
2281 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG

```

1 AGACACCTCT GCCCTCACCA TGAGCCTCTG GCAGCCCCTG GTCCTGGTGC TCCTGGTGCT
61 GGGCTGCTGC TTTGCTGCC CCAGACAGCG CCAGTCCACC CTTGTGCTCT TCCCTGGAGA
121 CCTGAGAAC AATCTCACCG ACAGGCAGCT GGCAGAGGAA TACCTGTACC GCTATGGTTA
181 CACTCGGGTG GCAGAGATGC GTGGAGAGTC GAAATCTCTG GGGCCTGCGC TGCTGCTTCT
241 CCAGAACCAA CTGTCCCTGC CCGAGACCGG TGAGCTGGAT AGCGCACGC TGAAGGCCAT
301 GCGAACCCCCA CGGTGCGGGG TCCCAGACCT GGGCAGATTG CAAACCTTTG AGGGCGACCT
361 CAAGTGGCAC CACCACAACA TCACCTATTG GATCCAAAAC TACTCGGAAG ACTTGCCGCG
421 GGCGGTGATT GACGACGCC TTGCCCCGCG CTTCGCACTG TGGAGCGCG TGACGCCGCT
481 CACCTTCACT CGCGTGTACA GCCGGGACGC AGACATCGTC ATCCAGTTTG GTGTCGCGGA
541 GCACGGAGAC GGGTATCCCT TCGACGGAA GGACGGCTC CTGGCACACG CCTTTCCCTC
601 TGGCCCCGGC ATTCAAGGGAG ACGCCCATT CGACGATGAC GAGTTGTGGT CCCTGGGCAA
661 GGGCGTCGTG GTTCCAACTC GGTTGGAAA CGCAGATGGC GCGGCCTGCC ACTTCCCCCTT
721 CATCTCGAG GGCGCTCCT ACTCTGCCTG CACCACCGAC GGTCGCTCCG ACGGGTTGCC
781 CTGGTGCAGT ACCACGGCCA ACTACGACAC CGACGACCGG TTTGGCTTCT GCCCCAGCGA
841 GAGACTCTAC ACCCGGGACG GCAATGCTGA TGGGAAACCC TGCCAGTTTC CATTCATCTT
901 CCAAGGCCAA TCCTACTCCG CCTGCACCAAC GGACGGTCGC TCCGACGGCT ACCGCTGGTG
961 CGCACCAACC GCCAACTACG ACCGGGACAA GCTCTCGGC TTCTGCCGA CCCGAGCTGA
1021 CTCGACGGTG ATGGGGGCA ACTCGGCGGG GGAGCTGTGC GTCTCCCTC TCAC TTTCCCT
1081 GGGTAAAGGAG TACTCGACCT GTACCAAGCGA GGGCCGCGGA GATGGGCGCC TCTGGTGC
1141 TAC ACCCTCG AACTTTGACA GCGACAAGAA GTGGGGCTTC TGCCCGGACC AAGGATACAG
1201 TTTGTTCCCTC GTGGCGGGCGC ATGAGTTCGG CCACGCGCTG GGCTTAGATC ATTCTCAGT
1261 GCGGGAGGCG CTCATGTACC CTATGTACCG CTTCACTGAG GGGCCCCCCT TGATAAGGA
1321 CGACGTGAAT GGCATCCGGC ACCTCTATGG TCCTCGCCCT GAACCTGAGC CACGGCCTCC
1381 AACCAACCAC ACACCGCAGC CCACGGCTCC CCCGACGGTC TGCCCCACCG GACCCCCCAC
1441 TGTCCACCCC TCAGAGCGCC CCACAGCTGG CCCCACAGGT CCCCCCTCAG CTGGCCCCAC
1501 AGGTCCCCCCC ACTGCTGGCC CTTCTACGGC CACTACTGTG CCTTTGAGTC CGGTGGACGA
1561 TGCCTGCAAC GTGAACATCT TCGACGCCAT CGCGGAGATT GGGAACAGC TGTATTGTT
1621 CAAGGATGGG AAGTACTGGC GATTCTCTGA GGGCAGGGGG AGCCGGCCGC AGGGCCCCCTT
1681 CCTTATCGCC GACAAGTGGC CGCGCTGCC CCCGAAGCTG GACTCGGTCT TTGAGGAGCC
1741 GCTCTCCAAG AAGCTTTCTC TCTTCTCTGG GCGCCAGGTG TGGGTGTACA CAGGCGCGTC
1801 GGTGCTGGGC CCGAGGCAGT TGGACAAGCT GGGCCTGGGA GCCGACGTGG CCCAGGTGAC
1861 CGGGGCCCTC CGGAGTGGCA GGGGAAAGAT GCTGCTGTT AGCGGGCGGC GCCTCTGGAG
1921 GTTCGACGTG AAGGCGCAGA TGGTGGATCC CCGGAGCGCC AGCGAGGTGG ACCGGATGTT
1981 CCCCGGGGTG CCTTTGGACA CGCACGACGT CTTCCAGTAC CGAGAGAAAG CCTATTTCTG
2041 CCAGGACCGC TTCTACTGGC GCGTGAGTTC CCGGAGTGAG TTGAACCAGG TGGACCAAGT
2101 GGGCTACGTG ACCTATGACA TCCTGCAGTG CCCTGAGGAC TAGGGCTCCC GTCTGCTTT
2161 GCAGTGCCAT GTAAATCCCC ACTGGGACCA ACCCTGGGGAG AGGAGCCAGT TTGCCGGATA
2221 CAA ACTGGTA TTCTGTTCTG GAGGAAAGGG AGGAGTGGAG GTGGGCTGGG CCCTCTCTTC
2281 TCAC CTTGTGTTCTA ATAAACTTGG ATTCTCTAAC CTTT

Figure 1

HUMAN OSTEOCLAST-SPECIFIC AND
-RELATED GENES

RELATED APPLICATION

This application is a continuation of application Ser. No. 08/045,270 filed on Apr. 6, 1993 now abandoned.

BACKGROUND OF THE INVENTION

Excessive bone resorption by osteoclasts contributes to the pathology of many human diseases including arthritis, osteoporosis, periodontitis, and hypercalcemia of malignancy. During resorption, osteoclasts remove both the mineral and organic components of bone (Blair, H. C., et al., *J. Cell Biol.* 102:1164 (1986)). The mineral phase is solubilized by acidification of the sub-osteoclastic lacuna, thus allowing dissolution of hydroxyapatite (Vaes, G., *Clin. Orthop. Relat.* 231:239 (1988))). However, the mechanism(s) by which type I collagen, the major structural protein of bone, is degraded remains controversial. In addition, the regulation of osteoclastic activity is only partly understood. The lack of information concerning osteoclast function is due in part to the fact that these cells are extremely difficult to isolate as pure populations in large numbers. Furthermore, there are no osteoclastic cell lines available. An approach to studying osteoclast function that permits the identification of heretofore unknown osteoclast-specific or -related genes and gene products would allow identification of genes and gene products that are involved in the resorption of bone and in the regulation of osteoclastic activity. Therefore, identification of osteoclast-specific or -related genes or gene products would prove useful in developing therapeutic strategies for the treatment of disorders involving aberrant bone resorption.

SUMMARY OF THE INVENTION

The present invention relates to isolated DNA sequences encoding all or a portion of osteoclast-specific or -related gene products. The present invention further relates to DNA constructs capable of replicating DNA encoding osteoclast-specific or -related gene products. In another embodiment, the invention relates to a DNA construct capable of directing expression of all or a portion of the osteoclast-specific or -related gene product in a host cell.

Also encompassed by the present invention are prokaryotic or eukaryotic cells transformed or transfected with a DNA construct encoding all or a portion of an osteoclast-specific or -related gene product. According to a particular embodiment, these cells are capable of replicating the DNA construct comprising the DNA encoding the osteoclast-specific or -related gene product, and, optionally, are capable of expressing the osteoclast-specific or -related gene product. Also claimed are antibodies raised against osteoclast-specific or -related gene products, or portions of these gene products.

The present invention further embraces a method of identifying osteoclast-specific or -related DNA sequences and DNA sequences identified in this manner. In one embodiment, cDNA encoding osteoclast is identified as follows: First, human giant cell tumor of the bone was used to 1) construct a cDNA library; 2) produce ³²P-labelled cDNA to use as a stromal cell⁺; osteoclast⁺ probe, and 3) produce (by culturing) a stromal cell population lacking osteoclasts. The presence of osteoclasts in the giant cell tumor was confirmed by histological staining for the osteoclast marker, type 5 acid phosphatase (TRAP). The cultured stromal cell population did not contain osteoclasts. The cultured stromal

clast marker, type 5 tartrate-resistant acid phosphatase (TRAP) and with the use of monoclonal antibody reagents.

The stromal cell population lacking osteoclasts was produced by dissociating cells of a giant cell tumor, then growing and passaging the cells in tissue culture until the cell population was homogeneous and appeared fibroblastic. The cultured stromal cell population did not contain osteoclasts. The cultured stromal cells were then used to produce a stromal cell⁺, osteoclast⁻ ³²P-labelled cDNA probe.

The cDNA library produced from the giant cell tumor of the bone was then screened in duplicate for hybridization to the cDNA probes: one screen was performed with the giant cell tumor cDNA probe (stromal cell⁺, osteoclast⁺), while a duplicate screen was performed using the cultured stromal cell cDNA probe (stromal cell⁺, osteoclast⁻). Hybridization to a stromal⁺, osteoclast⁺ probe, accompanied by failure to hybridize to a stromal⁺, osteoclast⁻ probe indicated that a clone contained nucleic acid sequences specifically expressed by osteoclasts.

In another embodiment, genomic DNA encoding osteoclast-specific or -related gene products is identified through known hybridization techniques or amplification techniques. In one embodiment, the present invention relates to a method of identifying DNA encoding an osteoclast-specific or -related protein, or gene product, by screening a cDNA library or a genomic DNA library with a DNA probe comprising one or more sequences selected from the group consisting of the DNA sequences set out in Table I (SEQ ID NOS: 1-32). Finally, the present invention relates to an osteoclast-specific or related protein encoded by a nucleotide sequence comprising a DNA sequence selected from the group consisting of the sequences set out in Table I, or their complementary strands.

35

BRIEF DESCRIPTION OF FIG. 1

The FIG. 1 shows cDNA sequence (SEQ ID NO: 33) of human gelatinase B, and highlights those portions of the sequence represented by the osteoclast-specific or -related cDNA clones of the present invention.

DETAILED DESCRIPTION OF THE
INVENTION

As described herein, Applicant has identified osteoclast-specific or osteoclast-related nucleic acid sequences. These sequences were identified as follows: Human giant cell tumor of the bone was used to 1) construct a cDNA library; 2) produce ³²P-labelled cDNA to use as a stromal cell⁺, osteoclast⁺ probe, and 3) produce (by culturing) a stromal cell population lacking osteoclasts. The presence of osteoclasts in the giant cell tumor was confirmed by histological staining for the osteoclast marker, type 5 acid phosphatase (TRAP). In addition, monoclonal antibody reagents were used to characterize the multinucleated cells in the giant cell tumor, which cells were found to have a phenotype distinct from macrophages and consistent with osteoclasts.

The stromal cell population lacking osteoclasts was produced by dissociating cells of a giant cell tumor, then growing the cells in tissue culture for at least five passages. After five passages the cultured cell population was homogeneous and appeared fibroblastic. The cultured population contained no multinucleated cells at this point, tested negative for type 5 acid phosphatase, and tested variably alkaline phosphatase positive. That is, the cultured stromal cell population did not contain osteoclasts. The cultured stromal

cells were then used to produce a stromal cell⁺, osteoclast-³²P-labelled cDNA probe.

The cDNA library produced from the giant cell tumor of the bone was then screened in duplicate for hybridization to the cDNA probes: one screen was performed with the giant cell tumor cDNA probe (stromal cell⁺, osteoclast⁺), while a duplicate screen was performed using the cultured stromal cell cDNA probe (stromal cell⁺ osteoclast⁻). Clones that hybridized to the giant cell tumor cDNA probe (stromal⁺, osteoclast⁺), but not to the stromal cell cDNA probe (stromal⁺, osteoclast⁻), were assumed to contain nucleic acid sequences specifically expressed by osteoclasts.

As a result of the differential screen described herein, DNA specifically expressed in osteoclast cells characterized as described herein was identified. This DNA, and equivalent DNA sequences, is referred to herein as osteoclast-specific or osteoclast-related DNA. Osteoclast-specific or -related DNA of the present invention can be obtained from sources in which it occurs in nature, can be produced recombinantly or synthesized chemically; it can be cDNA, genomic DNA, recombinantly-produced DNA or chemically-produced DNA. An equivalent DNA sequence is one which hybridizes, under standard hybridization conditions, to an osteoclast-specific or -related DNA identified as described herein or to a complement thereof.

Differential screening of a human osteoclastoma cDNA library was performed to identify genes specifically expressed in osteoclasts. Of 12,000 clones screened, 195 clones were identified which are either uniquely expressed in osteoclasts, or are osteoclast-related. These clones were further identified as osteoclast-specific, as evidenced by failure to hybridize to mRNA derived from a variety of unrelated human cell types, including epithelium, fibroblasts, lymphocytes, myelomonocytic cells, osteoblasts, and neuroblastoma cells. Of these, 32 clones contain novel cDNA sequences which were not found in the GenBank database.

A large number of cDNA clones obtained by this procedure were found to represent 92 kDa type IV collagenase (gelatinase B; E.C. 3.4.24.35) as well as tartrate resistant acid phosphatase. In situ hybridization localized mRNA for gelatinase B to multinucleated giant cells in human osteoclastomas. Gelatinase B immunoreactivity was demonstrated in giant cells from 8/8 osteoclastomas, osteoclasts in normal bone, and in osteoclasts of Paget's disease by use of a polyclonal antisera raised against a synthetic gelatinase B peptide. In contrast, no immunoreactivity for 72 kDa type IV collagenase (gelatinase A; E.C. 3.4.24.24), which is the product of a separate gene, was detected in osteoclastomas or normal osteoclasts.

The present invention has utility for the production and identification of nucleic acid probes useful for identifying osteoclast-specific or -related DNA. Osteoclast-specific or -related DNA of the present invention can be used to produce osteoclast-specific or -related gene products useful in the therapeutic treatment of disorders involving aberrant bone resorption. The osteoclast-specific or -related sequences are also useful for generating peptides which can then be used to produce antibodies useful for identifying osteoclast-specific or -related gene products, or for altering the activity of osteoclast-specific or -related gene products. Such antibodies are referred to as osteoclast-specific antibodies. Osteoclast-specific antibodies are also useful for identifying osteoclasts. Finally, osteoclast-specific or -related DNA sequences of the present invention are useful in gene therapy. For example, they can be used to alter the

expression in osteoclasts of an aberrant osteoclast-specific or -related gene product or to correct aberrant expression of an osteoclast-specific or -related gene product. The sequences described herein can further be used to cause osteoclast-specific or related gene expression in cells in which such expression does not ordinarily occur, i.e., in cells which are not osteoclasts.

Example 1—Osteoclast cDNA Library Construction

Messenger RNA (mRNA) obtained from a human osteoclastoma ('giant cell tumor of bone'), was used to construct an osteoclastoma cDNA library. Osteoclastomas are actively bone resorptive tumors, but are usually non-metastatic. In cryostat sections, osteoclastomas consist of ~30% multinucleated cells positive for tartrate resistant acid phosphatase (TRAP), a widely utilized phenotypic marker specific in vivo for osteoclasts (Minkin, *Calcif. Tissue Int.* 34:285-290 (1982)). The remaining cells are uncharacterized 'stromal' cells, a mixture of cell types with fibroblastic/mesenchymal morphology. Although it has not yet been definitively shown, it is generally held that the osteoclasts in these tumors are non-transformed, and are activated to resorb bone in vivo by substance(s) produced by the stromal cell element.

Monoclonal antibody reagents were used to partially characterize the surface phenotype of the multinucleated cells in the giant cell tumors of long bone. In frozen sections, all multinucleated cells expressed CD68, which has previously been reported to define an antigen specific for both osteoclasts and macrophages (Horton, M. A. and M. H. Helfrich, In *Biology and Physiology of the Osteoclast*, B. R. Rifkin and C. V. Gay, editors, CRC Press, Inc. Boca Raton, Fla., 33-54 (1992)). In contrast, no staining of giant cells was observed for CD11b or CD14 surface antigens, which are present on monocyte/macrophages and granulocytes (Arnaout, M. A. et al. *J. Cell. Physiol.* 137:305 (1988); Haziot, A. et al. *J. Immunol.* 141:547 (1988)). Cytocentrifuge preparations of human peripheral blood monocytes were positive for CD68, CD11b, and CD14. These results demonstrate that the multinucleated giant cells of osteoclastomas have a phenotype which is distinct from that of macrophages, and which is consistent with that of osteoclasts.

Osteoclastoma tissue was snap frozen in liquid nitrogen and used to prepare poly A⁺ mRNA according to standard methods. cDNA cloning into a pcDNAII vector was carried out using a commercially-available kit (Librarian, InVitrogen). Approximately 2.6×10⁶ clones were obtained, >95% of which contained inserts of an average length 0.6 kB.

Example 2—Stromal Cell mRNA Preparation

A portion of each osteoclastoma was snap frozen in liquid nitrogen for mRNA preparation. The remainder of the tumor was dissociated using brief trypsinization and mechanical disaggregation, and placed into tissue culture. These cells were expanded in Dulbecco's MEM (high glucose, Sigma) supplemented with 10% newborn calf serum (MA Bioproducts), gentamycin (0.5 mg/ml), L-glutamine (2 mM) and non-essential amino acids (0.1 mM) (Gibco). The stromal cell population was passaged at least five times, after which it showed a homogenous, fibroblastic looking cell population that contained no multinucleated cells. The stromal cells were mononuclear, tested negative acid phosphatase, and tested variably alkaline phosphatase positive. These findings indicate that propagated stromal cells (i.e., stromal cells that

are passaged in culture) are non-osteoclastic and non-activated.

Example 3—Identification of DNA Encoding Osteoclastoma-Specific or -Related Gene Products by Differential Screening of an Osteoclastoma cDNA Library

A total of 12,000 clones drawn from the osteoclastoma cDNA library were screened by differential hybridization, using mixed ^{32}P labelled cDNA probes derived from (1) giant cell tumor mRNA (stromal cell $^+$, OC $^+$), and (2) mRNA from stromal cells (stromal cell $^+$, OC $^-$) cultivated from the same tumor. The probes were labelled with ^{32}P dCTP by random priming to an activity of $\sim 10^9\text{ CPM}/\mu\text{g}$. Of these 12,000 clones, 195 gave a positive hybridization signal with giant cell (i.e., osteoclast and stromal cell) mRNA, but not with stromal cell mRNA. Additionally, these clones failed to hybridize to cDNA produced from mRNA derived from a variety of unrelated human cell types including epithelial cells, fibroblasts, lymphocytes, myelomonocytic cells, osteoblasts, and neuroblastoma cells. The failure of these clones to hybridize to cDNA produced from mRNA derived from other cell types supports the conclusion that these clones are either uniquely expressed in osteoclasts, or are osteoclast-related.

The osteoclast (OC) cDNA library was screened for differential hybridization to OC cDNA (stromal cell $^+$, OC $^+$) and stromal cell cDNA (stromal cell $^+$, OC $^-$) as follows:

NYTRAN filters (Schleicher & Schuell) were placed on agar plates containing growth medium and ampicillin. Individual bacterial colonies from the OC library were randomly picked and transferred, in triplicate, onto filters with pre-ruled grids and then onto a master agar plate. Up to 200 colonies were inoculated onto a single 90-mm filter/plate using these techniques. The plates were inverted and incubated at 37° C. until the bacterial inoculates had grown (on the filter) to a diameter of 0.5–1.0 mm.

The colonies were then lysed, and the DNA bound to the filters by first placing the filters on top of two pieces of Whatman 3 MM paper saturated with 0.5N NaOH for 5 minutes. The filters were neutralized by placing on two pieces of Whatman 3 MM paper saturated with 1M Tris-HCL, pH 8.0 for 3–5 minutes. Neutralization was followed by incubation on another set of Whatman 3 MM papers saturated with 1M Tris-HCL, pH 8.0/1.5M NaCl for 3–5 minutes. The filters were then washed briefly in 2 \times SSC.

DNA was immobilized on the filters by baking the filters at 80° C. for 30 minutes. Filters were best used immediately, but they could be stored for up to one week in a vacuum jar at room temperature.

Filters were prehybridized in 5–8 ml of hybridization solution per filter, for 2–4 hours in a heat sealable bag. An additional 2 ml of solution was added for each additional filter added to the hybridization bag. The hybridization

buffer consisted of 5 \times SSC, 5 \times Denhardt's solution, 1% SDS and 100 $\mu\text{g}/\text{ml}$ denatured heterologous DNA.

Prior to hybridization, labeled probe was denatured by heating in 1 \times SSC for 5 minutes at 100° C., then immediately chilled on ice. Denatured probe was added to the filters in hybridization solution, and the filters hybridized with continuous agitation for 12–20 hours at 65° C.

After hybridization, the filters were washed in 2 \times SSC/0.2% SDS at 50°–60° C. for 30 minutes, followed by washing in 0.2 \times SSC/0.2% SDS at 60° C. for 60 minutes.

The filters were then air dried and autoradiographed using an intensifying screen at –70° C. overnight.

Example 4—DNA Sequencing of Selected Clones

Clones reactive with the mixed tumor probe, but unreactive with the stromal cell probe, are expected to contain either osteoclast-related, or *in vivo* 'activated' stromal-cell-related gene products. One hundred and forty-four cDNA clones that hybridized to tumor cell cDNA, but not to stromal cell cDNA, were sequenced by the dideoxy chain termination method of Sanger et al. (Sanger F, et al. *Proc. Natl. Acad. Sci. USA* 74:5463 (1977)) using sequenase (US Biochemical). The DNASIS (Hitatchi) program was used to carry out sequence analysis and a homology search in the GenBank/EMBL database.

Fourteen of the 195 tumor $^+$ stromal $^-$ clones were identified as containing inserts with a sequence identical to the osteoclast marker, type 5 tartrate-resistant acid phosphatase (TRAP) (GenBank accession number J04430 M19534). The high representation of TRAP positive clones also indicates the effectiveness of the screening procedure in enriching for clones which contain osteoclast-specific or related cDNA sequences.

Interestingly, an even larger proportion of the tumor $^+$ stromal $^-$ clones (77/195; 39.5%) were identified as human gelatinase B (macrophage-derived gelatinase) (Wilhelm, S. M. *J. Biol. Chem.* 264:17213 (1989)), again indicating high expression of this enzyme by osteoclasts. Twenty-five of the gelatinase B clones were identified by dideoxy sequence analysis; all 25 showed 100% sequence homology to the published gelatinase B sequence (Genbank accession number J05070). The portions of the gelatinase B cDNA sequence covered by these clones is shown in the FIGURE (SEQ ID NO: 33). An additional 52 gelatinase B clones were identified by reactivity with a ^{32}P -labelled probe for gelatinase B.

Thirteen of the sequenced clones yielded no readable sequence. A DNASIS search of GenBank/EMBL databases revealed that, of the remaining 91 clones, 32 clones contain novel sequences which have not yet been reported in the databases or in the literature. These partial sequences are presented in Table I. Note that three of these sequences were repeats, indicating fairly frequent representation of mRNA related to this sequence. The repeat sequences are indicated by ^a, ^b superscripts (Clones 198B, 223B and 32C of Table I).

TABLE I

PARTIAL SEQUENCES OF 32 NOVEL OC-SPECIFIC OR -RELATED EXPRESSED GENES (cDNA CLONES)

34A (SEQ ID NO: 1) 1 GCAAATATCT 61 AATGTTTCTA 121 GTGATATTCT 4B (SEQ ID NO: 2) 1 GTGTCAACCT	AAGTTTATTG GGGTTTTTTT CTTGAAATAA GCATATCCTA	CTTGGATITC AGTTTGTITI ACCTATAATA AAAATGTCAA	TAGTGAGAGC TATTGAAAAAA GAAAATAGCA AATGCTGCAT	TGTTGAATT TTAATTATT GCAGACAACA CTGGTTAATG	GGTGATGTCA TATGCTATAG TCGGGGTAGG
--	--	--	---	--	--

TABLE I-continued

PARTIAL SEQUENCES OF 32 NOVEL OC-SPECIFIC OR -RELATED
EXPRESSED GENES (cDNA CLONES)

61	GGG					
12B	(SEQ ID NO: 3)					
1	CTTCCCTCTC	TTGCCCTCCCT	TTCCCCAAGCA	GAGGTGCTCA	CTCCATGGCC	ACCGCCACCA
61	CAGGCCACA	GGGAGTACTG	CCAGACTACT	GCTGATGTC	TCTTAAGGCC	CAGGGAGTCT
121	CAACCAAGCTG	GTGGTGAATG	CTGCCCTGGCA	CGGGACCCCC	CCC	
28B	(SEQ ID NO: 4)					
1	TTTATATTTG	AAATATATGT	ATTACATCCC	TAGAAAAAGA	ATCCCAAGGAT	TTTCCCTCTC
61	GTGTGTTTC	GTCTTGCTTC	TTCATGGTCC	ATGATGCCAG	CTGAGGTGT	CAGTACAATG
121	AAACCAAACT	GGGGGGATGG	AAGCAGATA	TTCTGCCATT	TTTCCAGGTC	TTT
37B	(SEQ ID NO: 5)					
1	GGCTGGACAT	GGGTGCCCTC	CACGTCCCTC	ATATCCCCAG	GCACACTCTG	GCCTCAGGTT
61	TTGCCCTGGC	CATGTCATCT	ACCTGGAGTG	GGCCCTCCCCC	TTCTTCAGCC	TTGAATCAA
121	AGCCACTTG	TTAGGCGAGG	ATTTCCAGA	CCACTCATCA	CATTAaaaaaa	TATTTGAAA
181	ACAAAAAAA	AAAAAAA				
55B	(SEQ ID NO: 6)					
1	TTGACAAAGC	TGTATTTC	CACCAATAAA	TAGTATATGG	TGATTGGGTT	TTCTATTTAT
61	AAGAGTGTG	GCTATTATAT	GGGGTATCAT	GTITGATGCTC	ATAAAATAGTT	CATATCTACT
121	TAATTTGCTC	TC				
60B	(SEQ ID NO: 7)					
1	GAAGAGAGTT	GTATGTACAA	CCCCAACAGG	CAAGGCAGCT	AAATGCAGAG	GGTACAGAGA
61	GATCCCGAGG	GAATT				
86B	(SEQ ID NO: 8)					
1	GGATGAAAC	ATGTAGAAGT	CCAGAGAAAA	ACAAATTTAA	AAAAGGTGG	AAAAGTTACG
61	GCAAAACCTGA	GATTTCAGCA	TAAAATCTT	AGTTAGAAGT	GAGAGAAAGA	AGAGGGAGGC
121	TGGTTGCTGT	TGCACTATC	AATAGTTAT	C		
87B	(SEQ ID NO: 9)					
1	TTCTTGATCT	TTAGAACACT	ATGAATAGGG	AAAAAAAGAAA	AAACTGTTCA	AAATAAAATG
61	TAGGAGCCGT	GCTTTGGAA	TGCTTGAGTG	AGGAGCTCAA	CAAGTCCTCT	CCCAAGAAAG
181	CAATGATAAA	ACTTGACAAA	A			
98B	(SEQ ID NO: 10)					
1	ACCCATTCT	AAACATTTT	ACTGTAAAT	TTTGGTCAA	AGTTCTAAC	TTAACACAT
61	CTAAAGAAT	AGAGGCAATA	TATAGCCAT	CTTACTAGAC	ATACAGTATT	AAACTGGACT
121	GAATATGAGG	ACAAGCTCTA	GTGGTCATTA	AACCCCTCAG	AA	
110B	(SEQ ID NO: 11)					
1	ACATATATTA	ACAGCATTCA	TTTGGCCAAA	ATCTACACGT	TTGAGAACATC	CTACTGTATA
61	TAAAGTGGGA	ATGTATCAAG	TATAGACTAT	GAAAGTGCAA	ATAACAAGTC	AAGGTTAGAT
121	TAACTTTTT	TTTTTACATT	ATAAAATTAA	CTTGT		
118B	(SEQ ID NO: 12)					
1	CCAAATTCT	CTGGAATCCA	TCCTCCCTCC	CATCACCATA	GCCTCGAGAC	GTCATTTCTG
61	TTTGACTACT	CCAGC				
133B	(SEQ ID NO: 13)					
1	AACTAACCTC	CTCGGACCCC	TGCCTCACTC	ATTACACCA	ACCACCAAC	TATCTATAAA
61	CCTGAGCCAT	GGCCATCCCT	TATGAGCGGC	GCAGTGATTA	TAGGCTTCG	CTCTAAGATA
121	AAAT					
140B	(SEQ ID NO: 14)					
1	ATATTATTTC	TTTTTTATG	TTAGCTTAGC	CATGAAAAT	TTACTGGTGA	AGCAGTTAAT
61	AAAACACACA	TCCCAATTGAA	GGTTTTTGTA	CATTTCAGTC	CTTACAAATA	ACAAAGCAAT
121	GATAAACCCG	GCACGTCTG	ATAGGAAATT	C		
144B	(SEQ ID NO: 15)					
1	CGTGACACAA	ACATGCATT	GTTTTATCA	TAAAACAGCC	TGGTTTCCTA	AAACAATACA
61	AACAGCATGT	TCATCAGCAG	GAAGCTGGCC	GTGGGCGAGG	GGGCC	
198B ^a	(SEQ ID NO: 16)					
1	ATAGGTTAGA	TTCTCATTCA	CGGGACTAGT	TAGCTTAAG	CACCCTAGAG	GACTAGGGTA
61	ATCTGACTTC	TCACCTCCCA	AGTCCCCCT	TATATCCTCA	AGGTAGAAAT	GTCTATGTTT
121	TCTACTCCAA	TTCATAAATC	TATTCTAAAG	TCTTGGTAC	AAGTTACATG	ATAAAAAGAA
181	ATGTGATTTG	TCTTCCCTTC	TTTGCACTT	TRAATAAAAG	TATTATCTC	CTGTCACAG
241	TTTAAT					
212B	(SEQ ID NO: 17)					
1	GTCCAGTATA	AAGGAAACGG	TTAAGTCGGT	AAGCTAGAGG	ATTGTAATA	TCTTTATGT
61	CCTCTAGATA	AAACACCCGA	TTAACAGATG	TTAACCTTT	ATGTTTGAT	TTGCTTTAAA
121	ATGCGCTTC	TACACATTAG	CTCCAGCTAA	AAAGACACAT	TGAGAGCTTA	GAGGAGATGC
181	TCTGGAGC					
223B ^b	(SEQ ID NO: 18)					
1	GCAGTTGGAA	GGGAGTTGGT	GTGCTATTTC	TGAAGCAGAT	GTGGTGATAC	TGAGATTGTC
61	TGTTCACTT	CCCCATTGTC	TTGTGCTTC	AAATGATCCTT	CCTACTTTGC	TTCTCTCCAC
121	CCATGACCTT	TTTCACTGTG	GCCATCAAGG	ACTTTCTGTA	CAGCTGTGT	ACTCTTAGGC
181	TAAGAGATGT	GACTACAGCC	TGCCCCCTGAC	TG		
241B	(SEQ ID NO: 19)					
1	TGTTAGTTT	TAGGAAGGCC	TGTCTTCTGG	GAGTGGAGTT	TATTAGTCCA	CTTCTGGAG
61	CTAGACGTCC	TATAGTTAGT	CACTGGGGAT	GGTGAAGAG	GGAGAAGAGG	AAGGGCGAAG
121	GGAGGGCTC	TTTGCTAGTA	TCTCCATTTC	TAGAAAGATGG	TTTATGATGAT	AACCACAGGT
181	CTATATGAGC	ATGTAAGGC	TGT			
32C ^b	(SEQ ID NO: 20)					
1	CCTATTTCTG	ATCCGTACTT	TGGACAAGGC	CCTTCAGCCA	GAAGACTGAC	AAAGTCATCC
121	TCCGTCTACC	AGAGCGTGCA	CTTGTGATCC	TAAGATAAGC	TTICATCTCG	GCTGTGCCTT
161	GGGTGGAAGG	GGCAGGATTC	TGCAGCTGCT	TTTGCTATTTC	TCTTCCTAAA	TTTCATT

TABLE I-continued

PARTIAL SEQUENCES OF 32 NOVEL OC-SPECIFIC OR -RELATED EXPRESSED GENES (cDNA CLONES)

34C (SEQ ID NO: 21)					
1 CGGAGCGTAG	GTGTGTAT	TCCGTACAA	ATCATTACAA	AACCAAGTCT	GGGGCAGTCA
61 CCGCCCCCAC	CCATCACCCC	AGTGAATGG	CTAGCTGCTG	GCCTTT	
47C (SEQ ID NO: 22)					
1 TTAGTTCACT	CAAAGCAGGC	AACCCCTTT	GGCACTGCTG	CCACTGGGGT	CATGGCGGT
61 GTGGCAGCTG	GGGAGGTTTC	CCCAACACCC	TCCCTGCTT	CCCTGTGT	CGGGGTCTCA
121 GGAGCTGACC	CAGAGTGG				
65C (SEQ ID NO: 23)					
1 GCTGAATGTT	TAAGAGAGAT	TTTGGTCTTA	AAGGCTCAT	CATGAAAGTG	TACATGCATA
61 TGCAAGTGTG	AATTACGTGG	TATGGATGGT	TGCTTGTGTTA	TTAACTAAAG	ATGTACAGCA
121 AACTGCCGT	TTAGAGTCCT	CTTAATATTG	ATGCTCTAAC	ACTGGGTCTG	CTTATGC
79C (SEQ ID NO: 24)					
1 GGCAGTGGGA	TATGGAATCC	AGAAGGGAAA	CAAGCACTGG	ATAATTAAAA	ACAGCTGGGG
61 AGAAAACCTGG	GGAAACAAAG	GATATATCCT	CATGGCTCGA	AATAAGAACAA	ACGCCCTGTGG
121 CATTGCCAAC	CTGGCCAGCT	TCCCCAAGAT	GTGACTCCAG	CCAGAAA	
84C (SEQ ID NO: 25)					
1 GCCAGGGCGG	ACCGCTTTA	TTCCCTCTCT	GCCTCAGAGG	TCAGGAAGGA	GGTCTGGCAG
61 GACCTGCACTG	GGGCCCTAGT	CATCTGTCGC	AGCGAAGGTG	AAGGGACTCA	CCTTGTGCGC
121 COTGCCCTGAG	TAGAACCTGT	TCTGGAATTC	C		
86C (SEQ ID NO: 26)					
1 AACTCTTCA	CACTCTGGTA	TTTTTGTGTT	AACAATATAT	GTGTTGTGTC	TTGGAATTAA
61 GTTCATATCA	ATTCAATATG	AGCTGTCTCA	TTCCTTTTTT	AATGGTCATA	TACAGTAGTA
121 TTCAATTATA	AGAATATATC	CTAATACCTT	TTAAAA		
87C (SEQ ID NO: 27)					
1 GGATAAGAAA	GAAGGCCCTGA	GGCCTAGGGG	CCGRGGCTGG	CCTGCCTCTC	AGTCCTGGGA
61 CGCAGCAGCC	CGCACAGTT	GAGAGGGGCA	CITCCTCTTG	CTTAGGTTGG	TGAGGATCTG
121 TCGCTTGTTG	GGCGGTGGAG	AGCCACAAAAA			
88C (SEQ ID NO: 28)					
1 CTGACCTTCG	AGAGTTTGAC	CTGGAGCCGG	ATACCTACTG	CCGCTATGAC	TCGGTCAGCG
61 TGTTCAACGG	AGCCGTGAGC	GACGACTCCG	GTGGGGAAGT	TCTGCAGCGA	T
89C (SEQ ID NO: 29)					
1 ATCCCTGGCT	GTGGATAGTG	CTTTTGTTA	GCAATGCTC	CCTCCTTAAG	GTTATAGGGC
61 TCCCTGAGTT	TGGGAGTGTG	GAAGTACTAC	TTAACCTGCT	GTCCCTGCTTG	GCTGTOGTTA
121 TCGTTTCTG	GTGATGTTGT	GCTAACAAATA	AGAATAC		
101C (SEQ ID NO: 30)					
1 GGCTGGGCAT	CCCTCTCCTC	CTCCATCCCC	ATACATCACC	AGGCTTAATG	TTTACAAACG
61 GTGCAAGGCC	GGCTCTGAAG	CCAAGGGCG	TCCGTGCCAC	GGTGGCTGTG	AGTATCCCTC
121 CGTTAGCTTT	CCCATAAAGT	TGGAGTATCT	GC		
112C (SEQ ID NO: 31)					
1 CCAACTCTA	CCCGATACA	GACCCACAGA	GTGCCATCCC	TGAGAGACCA	GACCGCTCCC
161 CAATACTCTC	CTAAAATAAA	CATGAAGCAC			
114C (SEQ ID NO: 32)					
1 CATGGATGAA	TGTCTCATGG	TGGGAAGGAA	CATGGTACAT	TTC	

^aRepeated 3 times^bRepeated 2 times

Sequence analysis of the OC⁺ stromal cell⁻ cloned DNA sequences revealed, in addition to the novel sequences, a number of previously-described genes. The known genes identified (including type 5 acid phosphatase, gelatinase B, cystatin C (13 clones), Alu repeat sequences (11 clones), creatinine kinase (6 clones) and others) are summarized in Table II. In situ hybridization (described below) directly demonstrated that gelatinase B mRNA is expressed in multinucleated osteoclasts and not in stromal cells. Although gelatinase B is a well-characterized protease, its expression at high levels in osteoclasts has not been previously described. The expression in osteoclasts of cystatin C, a cysteine protease inhibitor, is also unexpected. This finding has not yet been confirmed by in situ hybridization. Taken together, these results demonstrate that most of these identified genes are osteoclast-expressed, thereby confirming the effectiveness of the differential screening strategy for identifying DNA encoding osteoclast-specific or -related gene products. Therefore, novel genes identified by this method have a high probability of being OC-specific or related.

In addition, a minority of the genes identified by this screen are probably not expressed by OCs (Table II). For example, type III collagen (6 clones), collagen type I (1 clone), dermatansulfate (1 clone), and type VI collagen (1

clone) are more likely to originate from the stromal cells or from osteoblastic cells which are present in the tumor. These cDNA sequences survive the differential screening process either because the cells which produce them in the tumor in vivo die out during the stromal cell propagation phase, or because they stop producing their product in vitro. These clones do not constitute more than 5-10% of the all sequences selected by differential hybridization.

TABLE II

SEQUENCE ANALYSIS OF CLONES ENCODING KNOWN SEQUENCES FROM AN OSTEOCLASTOMA cDNA LIBRARY

Clones with Sequence Homology to Collagenase Type IV	25 total
Clones with Sequence Homology to Type 5 Taurate Resistant Acid Phosphatase	14 total
Clones with Sequence Homology to Cystatin C:	13 total
Clones with Sequence Homology to Alu-repeat Sequences	11 total
Clones with Sequence Homology to Creatinine Kinase	6 total
Clones with Sequence Homology to	6 total

TABLE II-continued

SEQUENCE ANALYSIS OF CLONES ENCODING KNOWN SEQUENCES FROM AN OSTEOCLASTOMA cDNA LIBRARY		5
Type III Collagen		
Clones with Sequence Homology to MHC Class I γ Invariant Chain	5 total	
Clones with Sequence Homology to MHC Class II β Chain	3 total	
One or Two Clone(s) with Sequence Homology to Each of the Following:	10 total	
$\alpha 1$ collagen type I		
γ interferon inducible protein		
osteopontin		
Human chondroitin/dermatansulfate		
α globin	15	
β glucosidase/sphingolipid activator		
Human CAPL protein (Ca binding)		
Human EST 01024		
Type VI collagen		
Human EST 00553		

12
UTP digoxigenin labelled cRNA probes.

TABLE III

5	In Situ HYBRIDIZATION USING PROBES DERIVED FROM NOVEL SEQUENCES	Reactivity with:	
		Clone	Osteoclasts
10	4B	+	+
	28B*	+	-
	37B	+	+
	86B	-	-
	87B	-	-
	88C	+	+
	98B	+	+
	118B*	+	-
	140B*	+	-
	198B*	+	-
	212B*	+	-
	Gelatinase B*	+	-

20 *OC-expressed, as indicated by reactivity with antisense probe and lack of reactivity with sense probe on OCs only.

Example 5—In situ Hybridization of OC-Expressed Genes

In situ hybridization was performed using probes derived from novel cloned sequences in order to determine whether the novel putative OC-specific or -related genes are differentially expressed in osteoclasts (and not expressed in the stromal cells) of human giant cell tumors. Initially, in situ hybridization was performed using antisense (positive) and sense (negative control) cRNA probes against human type IV collagenase/gelatinase B labelled with 35 S-UTP.

A thin section of human giant cell tumor reacted with the antisense probe resulted in intense labelling of all OCs, as indicated by the deposition of silver grains over these cells, but failed to label the stromal cell elements. In contrast, only minimal background labelling was observed with the sense (negative control) probe. This result confirmed that gelatinase B is expressed in human OCs.

In situ hybridization was then carried out using cRNA probes derived from 11/32 novel genes, labelled with digoxigenin UTP according to known methods.

The results of this analysis are summarized in Table III. Clones 28B, 118B, 140B, 198B, and 212B all gave positive reactions with OCs in frozen sections of a giant cell tumor, as did the positive control gelatinase B. These novel clones therefore are expressed in OCs and fulfill all criteria for OC-relatedness. 198B is repeated three times, indicating relatively high expression. Clones 4B, 37B, 88C and 98B produced positive reactions with the tumor tissue; however the signal was not well-localized to OCs. These clones are therefore not likely to be useful and are eliminated from further consideration. Clones 86B and 87B failed to give a positive reaction with any cell type, possibly indicating very low level expression. This group of clones could still be useful but may be difficult to study further. The results of this analysis show that 5/11 novel genes are expressed in OCs, indicating that ~50% of novel sequences likely to be OC-related.

To generate probes for the in situ hybridizations, cDNA derived from novel cloned osteoclast-specific or -related cDNA was subcloned into a BlueScript II SK(-) vector. The orientation of cloned inserts was determined by restriction analysis of subclones. The T7 and T3 promoters in the 65 BlueScriptII vector was used to generate 35 S-labelled (35 S-UTP 850 Ci/mmol, Amersham, Arlington Heights, Ill.), or

25 tissue was fixed in 4% paraformaldehyde and embedded in OCT (Miles Inc., Kankakee, Ill.). The sections were rehydrated, postfixed in 4% paraformaldehyde, washed, and pretreated with 10 mM DTT, 10 mM iodoacetamide, 10 mM N-ethylmaleimide and 0.1 triethanolamine-HCl. Prehybridization was done with 50% deionized formamide, 10 mM Tris-HCl, pH 7.0, 1x Denhardt's, 500 mg/ml tRNA, 80 mg/ml salmon sperm DNA, 0.3M NaCl, mM EDTA, and 100 mM DTT at 45° C. for 2 hours. Fresh hybridization solution containing 10% dextran sulfate and 1.5 ng/ml 35S-labelled or digoxigenin labelled RNA probe was applied after heat denaturation. Sections were coverslipped and then incubated in a moistened chamber at 45°–50° C. overnight. Hybridized sections were washed four times with 50% formamide, 2x SSC, containing 10 mM DTT and 0.5% Triton X-100 at 45° C. Sections were treated with RNase A and RNase T1 to digest single-stranded RNA, washed four times in 2x SSC/10 mM DTT.

30 In order to detect 35 S-labelling by autoradiography, slides were dehydrated, dried, and coated with Kodak NTB-2 emulsion. The duplicate slides were split, and each set was placed in a black box with desiccant, sealed, and incubated at 4° C. for 2 days. The slides were developed (4 minutes) and fixed (5 minutes) using Kodak developer D19 and Kodak fixer. Hematoxylin and eosin were used as counterstains.

35 In order to detect digoxigenin-labelled probes, a Nucleic Acid Detection Kit (Boehringer-Mannheim, Cat. #1175041) was used. Slides were washed in Buffer 1 consisting of 100 mM Tris/150 mM NaCl, pH 7.5, for 1 minute. 100 μ l Buffer 2 was added (made by adding 2 mg/ml blocking reagent as provided by the manufacturer) in Buffer 1 to each slide. The slides were placed on a shaker and gently swirled at 20° C.

40 Antibody solutions were diluted 1:100 with Buffer 2 (as provided by the manufacturer). 100 μ l of diluted antibody solution was applied to the slides and the slides were then incubated in a chamber for 1 hour at room temperature. The slides were monitored to avoid drying. After incubation with antibody solution, slides were washed in Buffer 1 for 10 minutes, then washed in Buffer 3 containing 2 mM levamisole for 2 minutes.

45 After washing, 100 μ l color solution was added to the slides. Color solution consisted of nitroblue/tetrazolium salt

13

(NBT) (1:225 dilution) 4.5 μ l, 5-bromo-4-chloro-3-indolyl phosphate (1:285 dilution) 3.5 μ l, levamisole 0.2 mg in Buffer 3 (as provided by the manufacturer) in a total volume of 1 ml. Color solution was prepared immediately before use.

After adding the color solution, the slides were placed in a dark, humidified chamber at 20° C. for 2–5 hours and monitored for color development. The color reaction was stopped by rinsing slides in TE Buffer.

The slides were stained for 60 seconds in 0.25% methyl green, washed with tap water, then mounted with water-based Permount (Fisher).

Example 6—Immunohistochemistry

Immunohistochemical staining was performed on frozen and paraffin embedded tissues as well as on cytocentrifuge preparations (see Table IV). The following antibodies were used: polyclonal rabbit anti-human gelatinase antibodies; Ab110 for gelatinase B; monoclonal mouse anti-human CD68 antibody (clone KP1) (DAKO, Denmark); Mol (anti-CD11b) and Mo2 (anti-CD14) derived from ATCC cell lines HB CRL 8026 and TIB 228/HB44. The anti-human gelatinase B antibody Ab110 was raised against a synthetic peptide with the amino acid sequence EALMYPMYRFTEGPPPLHK (SEQ ID NO: 34), which is specific for human gelatinase B (Corcoran, M. L. et al. *J. Biol. Chem.*, 267:515 (1992)).

Detection of the immunohistochemical staining was achieved by using a goat anti-rabbit glucose oxidase kit (Vector Laboratories, Burlingame Calif.) according to the manufacturer's directions. Briefly, the sections were rehydrated and pretested with either acetone or 0.1% trypsin. Normal goat serum was used to block nonspecific binding. Incubation with the primary antibody for 2 hours or overnight (Ab110:1/500 dilution) was followed by either a glucose oxidase labeled secondary anti-rabbit serum, or, in the case of the mouse monoclonal antibodies, were reacted with purified rabbit anti-mouse Ig before incubation with the secondary antibody.

Paraffin embedded and frozen sections from osteoclastomas (GCT) were reacted with a rabbit antiserum against gelatinase B (antibody 110) (Corcoran, M. L. et al. *J. Biol. Chem.* 267:515 (1992)), followed by color development with glucose oxidase linked reagents. The osteoclasts of a giant cell tumor were uniformly strongly positive for gelatinase B, whereas the stromal cells were unreactive. Control sections reacted with rabbit preimmune serum were negative. Identical findings were obtained for all 8 long bone giant cell tumors tested (Table IV). The osteoclasts present in three out of four central giant cell granulomas (GCG) of the mandible were also positive for gelatinase B expression. These neoplasms are similar but not identical to the long bone giant cell tumors, apart from their location in the jaws (Shafer, W. G. et al., *Textbook of Oral Pathology*, W. B. Saunders Company, Philadelphia, pp. 144–149 (1983)). In contrast, the multinucleated cells from a peripheral giant cell tumor, which is a generally non-resorptive tumor of oral soft tissue,

14

were unreactive with antibody (Shafer, W. G. et. al., *Textbook of Oral Pathology*, W. B. Saunders Company, Philadelphia, pp. 144–149 (1983)).

Antibody 110 was also utilized to assess the presence of gelatinase B in normal bone (n=3) and in Paget's disease, in which there is elevated bone remodeling and increased osteoclastic activity. Strong staining for gelatinase B was observed in osteoclasts both in normal bone (mandible of a 2 year old), and in Paget's disease. Staining was again absent in controls incubated with preimmune serum. Osteoblasts did not stain in any of the tissue sections, indicating that gelatinase B expression is limited to osteoclasts in bone. Finally, peripheral blood monocytes were also reactive with antibody 110 (Table IV).

TABLE IV

DISTRIBUTION OF GELATINASE B IN VARIOUS TISSUES		
	Samples	Antibodies tested Ab 110 gelatinase B
20	GCT frozen (n = 2)	
	giant cells	+
	stromal cells	-
25	GCT paraffin (n = 6)	
	giant cells	+
	stromal cells	-
30	central GCG (n = 4)	
	giant cells	+ (34)
	stromal cells	-
35	peripheral GCT (n = 4)	
	giant cells	-
	stromal cells	-
40	Paget's disease (n = 1)	
	osteoclasts	+
45	osteoblasts	-
	normal bone (n = 3)	
	osteoclasts	+
	osteoblasts	-
50	monocytes (cytocentrifuge)	+

Distribution of gelatinase B in multinucleated giant cells, osteoclasts, osteoblasts and stromal cells in various tissues. In general, paraffin embedded tissues were used for these experiments; exceptions are indicated.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments described herein. Such equivalents are intended to be encompassed by the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i i i) NUMBER OF SEQUENCES: 34

-continued

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 170 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCAAATATCT AAGTTTATTG CTTGGATTT	TAGTGAGAGC TGTGAATT	GGTGATGTCA	6 0
AATGTTCTA GGGTTTTTTT AGTTGTTT	TATTGAAAAAA TTAAATTATT	TATGCTATAG	1 2 0
GTGATATTCT CTTTGAATAA ACCTATAATA	GAAAATAGCA GCAGACAA		1 7 0

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 63 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GTGTCACACCT GCATATCCTA AAAATGTCAA AATGCTGCAT CTGGTTAATG	TCGGGGTAGG	6 0
GGG		6 3

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 163 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTTCCCTCTC TTGCTTCCCT TTCCCAAGCA GAGGTGCTCA CTCCATGGCC	ACCGCCACCA	6 0
CAGGCCACACA GGGAGTACTG CCAGACTACT GCTGATGTT	TCTTAAGGCC CAGGGAGTCT	1 2 0
CAACCAGCTG GTGGTGAATG CTGCCTGGCA CGGGACCCCC CCC		1 6 3

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 173 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TTTTATTTGT AAATATATGT ATTACATCCC TAGAAAAAGA ATCCAGGAT	TTTCCCTCCT	6 0
GTGTGTTTTC GTCTTGCTTC TTCATGGTCC ATGATGCCAG CTGAGGTTGT	CAGTACAATG	1 2 0
AAACCAAAACT GGCGGGATGG AAGCAGATTA TTCTGCCATT	TTTCCAGGTC	1 7 3

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 197 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double

5,552,281

17

18

-continued

(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGCTGGACAT GGGTGCCTC CACGTCCTC ATATCCCCAG GCACACTCTG GCCTCAGGTT	6 0
TTGCCCTGGC CATGTATCT ACCTGGAGTG GGCCCTCCCC TTCTTCAGCC TTGAATCAA	1 2 0
AGCCACTTTG TTAGGCAGGG ATTTCCCAGA CCACATCA CATTAAAAAA TATTTGAAA	1 8 0
ACAAAAAAA AAAAAAA	1 9 7

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 132 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TTGACAAAGC TGTTTATTC CACCAATAAA TAGTATATGG TGATTGGGT TTCTATTTAT	6 0
AAGAGTAGTG GCTATTATAT GGGGTATCAT GTTGATGCTC ATAAATAGTT CATATCTACT	1 2 0
TAATTTGCCT TC	1 3 2

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 75 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAAGAGAGTT GTATGTACAA CCCAACAGG CAAGGCAGCT AAATGCAGAG GGTACAGAGA	6 0
GATCCCGAGG GAATT	7 5

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 151 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGATGAAAC ATGTAGAAGT CCAGAGAAAA ACAATTTAA AAAAAGGTGG AAAAGTTACG	6 0
GCAAACCTGA GATTCAGCA TAAAATCTTT AGTTAGAAGT GAGAGAAAGA AGAGGGAGGC	1 2 0
TGGTTGCTGT TGCACGTATC AATAGGTTAT C	1 5 1

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 141 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

-continued

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TTCTTGATCT TTAGAACACT ATGAATAGGG AAAAAAGAAA AACTGTTCA AAATAAAATG	6 0
TAGGAGCCGT GCTTTGGAA TGCTTGAGTG AGGAGCTCAA CAAGTCCTCT CCCAAGAAAG	1 2 0
CAATGATAAA ACTTGACAAA A	1 4 1

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 162 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ACCCATTCTTCTTAAACAATTTTACTGTAAAAT TTTGGTCAA AGTTCTAACG TTAATCACAT	6 0
CTCAAAAGAAT AGAGGCAATA TATAGCCCCT CTTACTAGAC ATACAGTATT AAACCTGGACT	1 2 0
GAATATGAGG ACAAGCTCTA GTGGTCATTA AACCCCTCAG AA	1 6 2

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 157 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ACATATATTA ACAGCATTCA TTTGGCCAAA ATCTACACGT TTGTAGAACAT CTACTGTATA	6 0
TAAAGTGGGA ATGTATCAAG TATAGACTAT GAAAGTGCAA ATAACAAGTC AAGGTTAGAT	1 2 0
TAACCTTTTT TTTTACATT ATAAAAATTAA CTTGTTT	1 5 7

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 75 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCAAATTTCT CTGGAAATCCA TCCTCCCTCC CATCACCCATA GCCTCGAGAC GTCATTTCTG	6 0
TTTGACTACT CCAGC	7 5

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 124 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AACATAACCTC CTCGGACCCC TGCCCTCACTC ATTTACACCA ACCACCCAAAC TATCTATAAA	6 0
CCTGAGCCAT GGCCATCCCT TATGAGCGGC GCAGTOATTA TAGGCTTCTG CTCTAAGATA	1 2 0

-continued

A A A T

1 2 4

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 151 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ATTATTATTATTC	TTTTTTTATG	TTAGCTTAGC	CATGCAAAAT	TTACTGGTGA	AGCAGTTAAT	6 0
AAAACACACA	TCCCATTGAA	GGGTTTGTA	CATTCAGTC	CTTACAAATA	ACAAAGCAAT	1 2 0
GATAAAACCCG	GCACGTCCCTG	ATAGGAAATT	C			1 5 1

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 105 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CGTGACACAA	ACATGCATTC	TTTTTATTCA	AAAAACAGCC	TGGTTTCCTA	AAACAATACA	6 0
AACAGCATGT	TCATCAGCAC	GAAGCTGGCC	GTGGGCAGGG	GGGCC		1 0 5

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 246 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ATAGGTTAGA	TTCTCATTCA	CGGGACTAGT	TAGCTTTAAG	CACCCCTAGAG	GACTAGGGTA	6 0
ATCTGACTTC	TCACTTCTCA	AGTTCCCTCT	TATATCCTCA	AGGTAGAAAT	GTCTATGTTT	1 2 0
TCTACTCCAA	TTCATAAAATC	TATTCTATAAG	TCTTGGTAC	AAGTTACATG	ATAAAAAGAA	1 8 0
ATGTGATTTG	TCTTCCCTTC	TTTGCACTTT	TGAAATAAAG	TATTTATCTC	CTGTCTACAG	2 4 0
TTTAAT						2 4 6

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 188 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GTCCAGTATA	AAGGAAAGCG	TTAAGTCGGT	AAGCTAGAGG	ATTGTAATA	TCTTTATGT	6 0
CCTCTAGATA	AAACACCCGA	TTAACAGATG	TTAACCTTT	ATGTTTGAT	TTGCTTTAAA	1 2 0
AATGGCCTTC	TACACATTAG	CTCCAGCTAA	AAAGACACAT	TOAGAGCTTA	GAGGATAGTC	1 8 0

5,552,281

23

24

-continued

TCTGGAGC

188

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 212 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GCACTTGGAA	GGGAGTTGGT	GTGCTATTTT	TGAAGCAGAT	GTGOTGATAAC	TGAGATTGTC	6 0
TGTTCACTTT	CCCCATTTGT	TTGTGCTTCA	AATGATCCTT	CCTACTTTGC	TTCTCTCCAC	1 2 0
CCATGACCTT	TTTCACTGTG	GCCATCAAGG	ACTTTCTGA	CAGCTTGTGT	ACTCTTAGGC	1 8 0
TAAGAGATGT	GAATACAGCC	TGCCCCCTGAC	TG			2 1 2

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 203 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TGTTAGTTTT	TAGGAACGCC	TGTCTTCTGG	GAOTGAGGTT	TATTAGTCCA	CTTCTTGGAG	6 0
CTAGACGTCC	TATAGTTAGT	CACTGGGGAT	GGTGAAAGAG	GGAGAAGAGG	AAGGGCGAAG	1 2 0
GGAAGGGCTC	TTTGCTAGTA	TCTCCATTTC	TAGAAGATGG	TTTAGATGAT	AACCACAGGT	1 8 0
CTATATGAGC	ATAGTAAGGC	TGT				2 0 3

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 177 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CCTATTTCTG	ATCCCTGACTT	TGGACAAGGC	CCTTCAGCCA	GAAGACTGAC	AAAGTCATCC	6 0
TCCGTCTACC	AGAGCGTGCA	CTTGTGATCC	TAAAATAAGC	TTCATCTCCG	GCTGTGCCTT	1 2 0
GGGTGGAAGG	GGCAGGATTC	TGCAGCTGCT	TTTGCATTTC	TCTTCCTAAA	TTTCATT	1 7 7

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 106 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CGGAGCGTAG	GTGTGTTAT	TCCTGTACAA	ATCATTACAA	AACCAAGTCT	GGGGCAGTCA	6 0
CCGGCCCCAC	CCATCACCCC	AGTGCATGG	CTAGCTGCTG	GCCTTT		1 0 6

5,552,281

25

26

-continued

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 139 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TTAGTTCAGT CAAAGCAGGC AACCCCTTT GGCACTGCTG CCACGGGGT CATGGCGGTT	6 0
GTGGCAGCTG GGGAGGTTTC CCCAACACCC TCCTCTGCTT CCCTGTGTGT CGGGGTCTCA	1 2 0
GGAGCTGACC CAGAGTGGA	1 3 9

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 177 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GCTGAATGTT TAAGAGAGAT TTTGGTCTTA AAGGCTTCAT CATGAAAGTG TACATGCATA	6 0
TGCAAGTGTG ATTACGTGG TATGGATGGT TGCTTGTTA TTAACTAAAG ATGTACAGCA	1 2 0
AACTGCCCCTT TTAGAGTCCT CTTAATATTG ATGTCTAAC ACTGGGTCTG CTTATGC	1 7 7

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 167 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GGCAGTGGGA TATGGAATCC AGAAGGGAAA CAAGCACTGG ATAATTAAAA ACAGCTGGGG	6 0
AGAAAACCTGG GGAAACAAAG GATATATCCT CATGGCTCGA AATAAGAACCA ACGCCTGTGG	1 2 0
CATTGCCAAC CTGGCCAGCT TCCCCAAGAT GTGACTCCAG CCAGAAA	1 6 7

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 151 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GCCAGGGCGG ACCGTCTTTA TTCCTCTCCT GCCTCAGAGG TCAGGAAGGA GGTCTGGCAG	6 0
GACCTGCAGT GGGCCCTAGT CATCTGTGGC AGCGAAGGTG AAGGGACTCA CCTTGTGCC	1 2 0
CGTGCCTGAG TAGAACTTGT TCTGGAATT C	1 5 1

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 156 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AACTCTTTCA CACTCTGGTA TTTTAGTTT AACAAATATAT GTGTTGTGTC TTGGAAATTAA	6 0
GTTCATATCA ATTCAATTG AGCTGTCTCA TTCTTTTTT AATGGTCATA TACAGTAGTA	1 2 0
TTCATTATA AGAATATATC CTAATACTTT TTAAAAA	1 5 6

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 150 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GGATAAGAAA GAAGGCCCTGA GGGCTAGGGG CGGGGGCTGG CCTGCCTCTC AGTCCTGGGA	6 0
CGCAGCAGCC CGCACAGGTT GAGAGGGGCA CTTCCTCTTG CTTAGGTTGG TGAGGATCTG	1 2 0
GTCCTGGTTG GCCGGTGGAG AGCCACAAAA	1 5 0

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 212 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GCACCTGGAA GGGAGTTGGT GTGCTATTT TGAAGCAGAT GTGGTGATAC TGAGATTGTC	6 0
TGTTCAAGTTT CCCCATTGTT TTGTGCTTCA AATGATCCTT CCTACTTGC TTCTCTCCAC	1 2 0
CCATGACCTT TTTCACTGTG GCCATCAAGG ACTTTCCCTGA CAGCTGTGT ACTCTTAGGC	1 8 0
TAAGAGATGT GACTACAGCC TGCCCCGTGAC TG	2 1 2

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 157 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:29:

ATCCCTGGCT GTGGATAGTG CTTTGTGTA GCAAATGCTC CCTCCTTAAG GTTATAGGGC	6 0
TCCCTGAGTT TGGGAGTGTG GAAGTACTAC TTAACTGTCT GTCCCTGCTTG GCTGTCGTTA	1 2 0
TCGTTTCTG GTGATGTTGT GCTAAACAATA AGAATAC	1 5 7

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 152 base pairs
- (B) TYPE: nucleic acid

-continued

(C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GGCTGGGCAT CCCTCTCCTC CTCCATCCCC ATACATCACC AGGTCTAATG TTTACAAACG	6 0
GTGCCAGCCC GGCTCTGAAG CCAAGGGCCG TCCGTGCCAC GGTGGCTGTG AGTATTCCCTC	1 2 0
CGTTAGCTTT CCCATAAGGT TGGAGTATCT GC	1 5 2

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 90 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CCAACTCCTA CCGCGATACA GACCCACAGA GTGCCATCCC TGAGAGACCA GACCGCTCCC	6 0
CAATACTCTC CTAAAAATAAA CATGAAGCAC	9 0

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CATGGATGAA TGTCTCATGG TGGGAAGGAA CATGGTACAT TTC	4 3
---	-----

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2333 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AGACACCTCT GCCCTCACCA TGAGCCTCTG GCAGCCCCTG GTCCCTGGTGC TCCTGGTGCT	6 0
GGGCTGCTGC TTTGCTGCC CCAGACAGCG CCAGTCCACC CTTGTGCTCT TCCCTGGAGA	1 2 0
CCTGAGAACCC AATCTCACCG ACAGGGCAGCT GGCAGAGGAA TACCTGTACC GCTATGGTTA	1 8 0
CACTCGGGTG GCAGAGATGC GTGGAGAGTC GAAAATCTCTG GGGCCTGCAC TGCTGCTTCT	2 4 0
CCAGAAGCAA CTGTCCCTGC CCGAGACCGG TGAGCTGGAT AGCGCCACGC TGAAGGCCAT	3 0 0
GCGAACCCCA CGGTGCGGGG TCCCAGACCT GGGCAGATTG CAAACCTTTG AGGGCGACCT	3 6 0
CAAGTGGCAC CACCCACAACA TCACCTATTG QATCCAAAAC TACTCGGAAG ACTTGCCCGCG	4 2 0
GGCGGTGATT GACGACGCCT TTGCCCCGCG CTTCGCACTG TGGAGCGCGG TGACGCCOCT	4 8 0
CACCTTCACT CGCGTGTACA GCGGGGACGC AGACATCGTC ATCCAGTTTG GTGTCGCGGA	5 4 0
GCACCGGAGAC GGGTATCCCT TCGACGGGAA GGACGGGCTC CTGGCACACCG CCTTTCCCTCC	6 0 0
TGGCCCCGGC ATTCAAGGGAG ACGCCCATTT CGACGATGAC GAGTTGTGGT CCCTGGGCAA	6 6 0

-continued

GGGC GTC GGT G	T T C C A A C T C	G G T T G G A A A	C G C A G A T G G C	G C G G C T G C C	A C T T C C C C T T	7 2 0
C A T C T T C G A G	G G C C G C T C C T	A C T C T G C C T G	C A C C A C C G A C	G G T C G C T C C G	A C G G G T T G C C	7 8 0
C T G G T G C A G T	A C C A C G G C C A	A C T A C G A C A C	C G A C G A C C G G	T T T G G C T T C T	G C C C C A G C G A	8 4 0
G A G A C T C T A C	A C C C G G G A C G	G C A A T G C T G A	T G G G A A A C C C	T G C C A G T T T C	C A T T C A T C T T	9 0 0
C C A A G G C C A A	T C C T A C T C C G	C C T G C A C C A C	G G A C G G T C G C	T C C G A C G G G C T	A C C G C T G G T G	9 6 0
C G C C A C C A C C	G C C A A C T A C G	A C C G G G A C A A	G C T C T T C G G C	T T C T G C C C G A	C C C G A G C T G A	1 0 2 0
C T C G A C G G T G	A T G G G G G G C A	A C T C G G G C G G G	G G A G C T G T G C	G T C T T C C C C T	T C A C T T T C C T	1 0 8 0
G G G T A A G G A G	T A C T C G A C C T	G T A C C A G C G A	G G G C C G C G G A	G A T G G G C G C C	T C T G G T G C G C	1 1 4 0
T A C C A C C T C G	A A C T T T G A C A	G C G A C A A G G A A	G T G G G G C T T C	T G C C C G G A C C	A A G G A T A C A G	1 2 0 0
T T T G T T C C T C	G T G G C G G C G C	A T G A G T T C G G	C C A C G C G C T G	G G C T T A G A T C	A T T C C T C A G T	1 2 6 0
G C C G G A G G C G	C T C A T G T A C C	C T A T G T A C C G	C T T C A C T G A G	G G G C C C C C C T	T G C A T A A G G A	1 3 2 0
C G A C G T G A A T	G G C A T C C G G C	A C C T C T A T G G	T C C T C G C C C T	G A A C C T G A G C	C A C G G C C T C C	1 3 8 0
A A C C A C C A C C	A C A C C G C A G C	C C A C G G C T C C	C C C G A C G G T C	T G C C C C A C C G	G A C C C C C A C	1 4 4 0
T G T C C A C C C C	T C A G A G G C G C C	C C A C A G C T G G	C C C C A C A G G T	C C C C C T C A G	C T G G C C C C A C	1 5 0 0
A G G T C C C C C C	A C T G C T G G C C	C T T C T A C G G C	C A C T A C T G T G	C C T T T G A G T C	C G G T G G A C G A	1 5 6 0
T G C C T G C A A C	G T G A A C A T C T	T C G A C G C C A T	C G C G G A G A T T	G G G A A C C A G C	T G T A T T T G T T	1 6 2 0
C A A G G A T G G G	A A G T A C T G G C	G A T T C T C T G A	G G G C A G G G G G	A G C C G G C C G C	A G G G C C C C T T	1 6 8 0
C C T T A T C G C C	G A C A A G T G G C	C C G C G C T G C C	C C G C A A G C T G	G A C T C G G T C T	T T G A G G A G C C	1 7 4 0
G C T C T C C A A G	A A G C T T T C T	T C T T C T C T G G	G C G C C A G G T G	T G G G T G T A C A	C A G G C G C G T C	1 8 0 0
G G T G C T G G G C	C C G A G G C G T C	T G G A C A A G C T	G G G C C T G G G A	G C C G A C G T O G	C C C A G G T G A C	1 8 6 0
C G G G G C C C T C	C G G A G T G G C A	G G G G G A A G A T	G C T G C T G T T C	A G C G G G C G G C	G C C T C T G G A G	1 9 2 0
G T T C G A C G T G	A A G G C G C A G A	T G G T G G A T C C	C C G G A G C G C C	A G C G A G G T G G	A C C G G A T G T T	1 9 8 0
C C C C G G G G T G	C C T T T G G A C A	C G C A C G A C G T	C T T C C A G T A C	C G A G A G A A A G	C C T A T T T C T G	2 0 4 0
C C A G G A C C G C	T T C T A C T G G C	G C G T G A G T T C	C C G G A G T G A G	T T G A A C C A G G	T G G A C C A A G T	2 1 0 0
G G G C T A C G T G	A C C T A T G A C A	T C C T G C A G T G	C C C T G A G G A C	T A G G G C T C C C	G T C C T G C T T	2 1 6 0
G C A G T G C C A T	G T A A A T C C C C	A C T G G G A C C A	A C C C T G G G G A	A G G A G C C A G T	T T G C C G G A T A	2 2 2 0
C A A A C T G G T A	T T C T G T T C T G	G A G G A A A G G G	A G G A G T G G A G	G T G G G C T G G G	C C C T C T C T C	2 2 8 0
T C A C C T T T G T	T T T T G T T G G	A G T G T T T C T A	A T A A A C T T G G	A T T C T C T A A C	C T T T	2 3 3 4

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:34:

G l u	A l a	L e u	M e t	T y r	P r o	M e t	T y r	A r g	P h e	T h r	G l u	G l y	P r o	P r o	L e u	
1																1 5
S																
1 0																
H i s L y s																

We claim:

1. An isolated osteoclast-specific or -related DNA sequence, or its complementary sequence, the DNA 65 sequence comprising a nucleic acid sequence selected from the group consisting of:

- a) DNA sequences set forth in the group consisting of SEQ ID NOS. 12, 14, 16 and 17, or their complementary strands; and

33

- b) DNA sequences which hybridize under standard conditions to the DNA sequences defined in a).
2. A DNA construct capable of replicating, in a host cell, osteoclast-specific or -related DNA, said construct comprising:
- 5
- a) a DNA sequence of claim 1; and
- b) sequences, in addition to said DNA sequence, necessary for transforming or transfecting a host cell, and for replicating, in a host cell, said DNA sequence.
3. A DNA construct capable of replicating and expressing, in a host cell, osteoclast-specific or -related DNA, said construct comprising:
- 10

34

- a) a DNA sequence of claim 2; and
- b) sequences, in addition to said DNA sequence, necessary for transforming or transfecting a host cell, and for replicating and expressing, in a host cell, said DNA sequence.
4. A cell stably transformed or transfected with a DNA construct according to claim 3.
5. A cell stably transformed or transfected with a DNA construct according to claim 4.

* * * * *